

The Prompt Engineering Fallacy: How Diverse Users Construct Prompts in Health Chatbot Interactions

Jia Lin^{1,2}

¹Department of Family, McGill University, Montreal, Canada

²Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montreal, Canada
jia.l.lin@mail.mcgill.ca

Abstract

Through a participatory health chatbot Promptathon, I reconceptualize prompting as a socially situated practice and introduce the prompt engineering fallacy.

engineering alone. I argue that this fallacy risks shifting responsibility for safe chatbot use onto users and may reproduce existing inequities in digital health literacy.

Background

Large language model (LLM) chatbots are increasingly used for health information seeking and decision-making (Clark and Bailey 2024). These systems can generate contextually relevant responses to a range of queries, making them attractive tools for expanding access to health information. Yet, they also raise significant concerns for safety, especially in healthcare contexts where users may interact with chatbots under stress, with limited clinical or technical knowledge, and without adequate oversight (Chow and Li 2025).

Prompt engineering has emerged as a strategy for improving chatbot performance. In healthcare, it has been used to improve clinical reasoning, medical question answering, diagnostic support, and patient-provider communication (Verawati and Yaqin 2025). Some scholars have framed prompt engineering as a key competency for effective human-chatbot interaction (Chen and Granet 2024). However, this framing often treats prompting as a technical task, assuming users can define tasks, specify outputs, and evaluate responses systematically.

Drawing on the social construction of technology (Pinch 2012), I explore health chatbot prompting as a socially situated practice. Through a participatory Promptathon, this study examined how participants constructed prompts to inform the design of a public-facing health chatbot.

This study makes three key contributions. First, it provides empirical evidence on diverse prompting practices. Second, it reconceptualizes prompting as a socially situated practice rather than a purely technical optimization task. Third, it introduces the prompt engineering fallacy, that safe chatbot use can be achieved primarily through user prompt

Methods

This study is part of a larger project co-designing a health-specific chatbot for families managing acute pediatric infectious symptoms (Lin et al. 2026). A Promptathon was held in 2026 to engage broader user perspectives in chatbot design. Twenty-three participants, including patients, public members, clinicians, health professionals, researchers, technical experts, and students, formed four interdisciplinary teams and interacted with three chatbot systems: a health-specific chatbot, GPT-5.3, and Gemini 3.1 Pro.

The Promptathon followed an iterative structure organized around two clinical scenarios: a pediatric vaccination case involving a sick toddler, and a heart failure case involving medication management and online misinformation. Participants first prompted individually and then reconvened with their teams to compare prompts and chatbot responses, refine approaches, and reflect on the quality and safety of the conversations.

Following rapid ethnographic approaches (Millen 2000), I treated the Promptathon as a concentrated field site for examining prompt design in practice. I observed activities and documented group interactions. Data collection included field notes, prompt artifacts, and audio-recorded discussions. I analyzed these materials using inductive and deductive thematic analysis to examine how participants constructed prompts across models and scenarios.

Results

My findings reconceptualize prompting as a socially situated practice shaped by participants' roles, clinical and technical knowledge, lived experience, and understandings of

chatbot function. Participants rarely approached prompting as a technical optimization task. Instead, prompting varied across participant groups and often remained stable across models and scenarios. Below, I highlight four themes.

First, **different participant groups enacted distinct prompting logics**. Patients and public members often used search-like or conversational prompts. They tended to provide limited contextual information and appeared to expect the chatbot to infer what was needed or ask follow-up questions. Clinicians, health professionals, and researchers tended to prioritize clinical completeness. Their prompts were longer, included more symptoms, and often requested evidence-based sources. Technical experts were more procedural, with some using explicit prompt engineering strategies, such as assigning roles, organizing prompts into sections, or reusing prompt templates. These differences reflected diverse expertise, as well as understandings of health chatbots, as a clinician, an assistant, or a configurable tool.

Second, **prompting practices often remained relatively stable**. Despite opportunities for refinement, participants often reused similar prompt across chatbots and clinical scenarios. When prompts changed, they generally became longer rather than more strategically structured. Notably, patient and public members were especially likely to maintain their conversational prompting style, sometimes resisting prompts that felt overly engineered as they did not reflect naturalistic health information seeking.

Third, **participants assessed chatbot conversations through different and sometimes conflicting criteria**. Team discussions revealed disagreement about what constituted a good prompt or conversation. Some participants prioritized accuracy, escalation advice, and evidence-based reasoning, while others emphasized clarity, reassurance, empathy, and situational appropriateness. Participants also tended to evaluate chatbot responses more directly than prompts themselves, attributing variation in responses to the chatbot than to the prompt.

Fourth, **safety was often treated as the chatbot's responsibility**. Participants rarely included safety instructions in their prompts and generally expected the chatbot to manage safety and risks. Safety became salient only when chatbot responses were inconsistent or unexpected, leading participants to become aware of chatbot limitations. Overall, safety was distributed across prompts, chatbot responses, and interpretations of risk.

Discussion

A central implication of this study is what we propose as the prompt engineering fallacy, the assumption that safe chatbot use can be achieved primarily through user prompt engineering alone. Our findings suggest that users, especially patients and public members, may not naturally approach

prompting in this way. Even when given opportunities to refine prompts, participants often added information rather than adopting formal prompt engineering strategies.

This mismatch between optimized prompting and naturalistic prompting has important ethical implications. A focus on prompt engineering may unintentionally risk shifting responsibility for interpreting and acting on chatbot responses onto users. In high-stakes health contexts, this may disadvantage those with lower digital health literacy, clinical knowledge, or prior experience with AI systems. Rather than democratizing access to health information, a reliance on prompt engineering may reproduce existing inequities.

These findings point to the need for health chatbots that scaffold naturalistic prompting. Rather than treating the first prompt as the main determinant for response quality, health chatbots should support naturalistic help seeking across interactions. Design implications include eliciting missing information, communicating uncertainty and limitations, and adapting to different prompting orientations. Such design could shift responsibility toward interactional support built into the chatbot itself.

Acknowledgements

This work is supported by a Canadian Institutes of Health Research Canada Graduate Scholarship Doctoral award.

References

- Chen, J. S.; and Granet, D.B. 2024. Prompt Engineering: Helping ChatGPT Respond Better to Patients and Parents. *Journal of Pediatric Ophthalmology & Strabismus* 61 (2): 148-149. doi.org/10.3928/01913913-20240124-02
- Chow, J. C. L.; and Li, K. 2025. Large Language Models in Medical Chatbots: Opportunities, Challenges, and The Need to Address AI Risks. *Information* 16 (7): 549. doi.org/10.3390/info16070549
- Clark, M.; and Bailey, S. 2024. Chatbots in Health Care: Connecting Patients to Information. *Canadian Journal of Health Technologies* 4 (1) 4 (1): 1-21.
- Lin, J.; Jaiswal, N.; Ma, Y.; Lebouché, B.; Villanueva, S.; Achiche, S. et al. 2026. Coconstructing CHAMP, an Artificial Intelligence Chatbot for Pediatric Infectious Symptoms Management: Protocol for a Multi-Phase Participatory Study. *JMIR Research Protocols*.
- Millen, D. R. 2000. Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*. doi.org/10.1145/347642.34776
- Pinch, T. 2012. The Social Construction of Technology: A Review. In *Technological Change*, 17-35.
- Verawati, N. N. S. P.; and Yaqin, L. N. 2025. Charting the Future of Prompt Engineering: Critical Reflections on Methodology, Ethics, and Research Directions. *International Journal* 4(1): 1. doi.org/10.36312/ijece.v4i1.2908